昆 虫 学 报 *Acta Entomologica Sinica*, May 2012, 55(5): 596 - 605 ISSN 0454-6296

蛾翅数学形态特征用于夜蛾分类和鉴定的可行性研究

蔡小娜[1,2]，黄大庄[1,*]，沈佐锐[2,*]，高灵旺[2]

(1. 河北农业大学林学院，河北保定 071000；2. 中国农业大学 IPMist 实验室，北京 100193)

摘要：为探讨蛾翅数学形态特征(MMC)在夜蛾科分类鉴定中的可行性，本文利用数字化技术获得和处理昆虫图像，对鳞翅目夜蛾科 6 种夜蛾的右前翅提取矩形度、延长度、叶状性、偏心率、球状性、似圆度和不变矩 Hu1、Hu2 等 13 项与大小尺度和方向均无关的数学形态特征，并利用方差分析、逐步判别分析和聚类分析等方法研究了各项数学形态特征在昆虫分类上作为分类特征的可行性、可靠性和重要性，并且从数学形态学角度对夜蛾科 6 个种的亲缘关系进行了分析。分析结果认为矩形度和延长度 2 个形态特征对这 6 种夜蛾的分类鉴定没有显著意义，从而筛选出 11 个形态特征作为分类变量，它们的作用大小依次为：(偏心率、Hu5、Hu7) > Hu2 > 似圆度 > 球状性 > Hu3 > (叶状性、Hu1、Hu6) > Hu4。利用蛾翅的这些特征参数成功地实现了对夜蛾科 6 种夜蛾的分类鉴定，基于这些特征参数的 6 种夜蛾的亲缘关系远近与基于传统形态学的系统进化观点相同。研究表明蛾翅数学形态特征可应用于蛾类昆虫的快速鉴定，为未来逐步实现蛾类昆虫的自动识别奠定了基础。

关键字：夜蛾；数学形态特征；翅；数字鉴定；特征提取；亲缘关系

中图分类号：Q964 **文献标识码**：A **文章编号**：0454-6296(2012)05-0596-10

Feasibility of math-morphological characters of wings for classification and identification of noctuid moths (Lepidoptera: Noctuidae)

CAI Xiao-Na[1,2], HUANG Da-Zhuang[1,*], SHEN Zuo-Rui[2,*], GAO Ling-Wang[2] (1. College of Forestry, Agricultural University of Hebei, Baoding, Hebei 071000, China; 2. IPMist Laboratory, China Agricultural University, Beijing 100193, China)

Abstract: To attest the feasibility of math-morphological characters (MMCs) of moth wings for classification and identification of Noctuidae, we used digital technology to acquire and process insect images. Thirteen MMCs, such as rectangularity, elongation, lobation, eccentricity, sphericity, roundness and movement invariants including Hu1 and Hu2, which are invariants to the image size and direction, were selected. The characters from right forewings of moths from six noctuid species were extracted in this study. The methods including variance analysis, stepwise discriminant analysis and cluster analysis were used to study the feasibility, reliability and importance of each MMC as classification attribute of insect classification. The kinship among the six noctuids was analyzed from the perspective of mathematical morphology. The analysis results showed that rectangularity and elongation had no significant contribution to classification and identification of the six noctuids. Therefore, 11 morphological characters were selected as the classification variables. The contribution of these variables could be ranked as follows: (eccentricity, Hu5, Hu7) > Hu2 > roundness > sphericity > Hu3 > (lobation, Hu1, Hu6) > Hu4. Classification and discrimination of the six noctuids were realized by using the character parameters of moth wings. The phylogenetic relationship of the six noctuids based on the character parameters was consistent with that based on the traditional morphological observation. The results indicate that MMCs of moth wings can be used to identify moths rapidly and will be useful in gradually realizing the automatic recognition of moths in the future.

Key words: Noctuid moths; math-morphological character (MMC); wings; digital identification; character extraction; phylogenetic relationship

基金项目：河北省自然科学基金项目(C2012204008)

作者简介：蔡小娜，女，1983 年 2 月生，河北保定人，博士研究生，研究方向为有害生物综合管理，E-mail：xuecaixiaona@163.com

* 通讯作者 Corresponding authors, E-mail：huangdazhuang@126.com；ipmist@163.com

收稿日期 Received：2012-02-28；接受日期 Accepted：2012-05-03

随着计算机技术的不断发展，数字图像处理和识别技术在人脸识别系统、车牌号码识别系统等人工智能领域得到广泛应用（张方樱等，2005；连建兵和王浩全，2011），而在生物学研究领域尤其是生物分类鉴定方面也日益起到重要的作用。英国自然历史博物馆的 MacLeod（2007）在其主编的《分类学中自动分类鉴定的理论方法和应用》一书中系统地介绍了生物学种类自动鉴定使用的理论方法以及 ABIS，DAISY 和 SPIDA 三大软件系统。O'Neill（2008）对 DAISY 软件在鳞翅目、双翅目、膜翅目以及其他节肢动物上的使用效果作了介绍，误判率约为 5%。Prieto 等（2009）运用形态测量学方法，通过对雄性外生殖器和翅形特征做主成分分析、判别分析以及聚类分析来探讨鳞翅目（Lepidoptera）灰蝶科（Lycaenidae）中 *Cupido minimus* 和 *C. carswelli* 是否为不同种的问题。在国内，中国农业大学 IPMist 实验室在这方面做了大量工作，首先，相继开发出了害虫鉴定辅导系统 PQ-InforMIS（王之岭，1996）以及昆虫自动鉴定软件 Bugvisux1.0（于新文，1999）和 BugShape1.0（张建伟，2006）；其次，应用 Bugvisux 系统，赵汗青等（2002，2003）对半翅目、鳞翅目和鞘翅目等昆虫整体图像提取了面积、周长、形状参数、叶状性、球状性、亮斑数等 11 项数学形态特征（math-morphological character，MMC）参数，探讨了根据昆虫的数学形态特征对昆虫实现自动鉴别的可行性；刘芳等（2008）利用数学形态学的彩色图像边缘检测算法在昆虫分类鉴定中进行了研究，潘鹏亮等（2008a，2008b）利用 DrawWing 和 TPSDig 翅脉特征点提取软件，分析了翅脉的数学形态特征在蝴蝶分类鉴定中的应用问题。这些研究结果都表明数学形态学对生物分类鉴定具有实际应用价值。

本文报道 IPMist 实验室的最新进展，即对蛾类进行数字化分类鉴定研究，一方面为了探讨上述成果在更多种类昆虫分类鉴定中应用的效果，另一方面力图在涵盖较多昆虫种类的情况下为数学形态学理论和方法的创新打下基础。蛾类有许多种类为农林重要害虫，其中夜蛾科绝大部分种类均为害虫，这对农林业造成了重要的危害和经济损失。当前对蛾类害虫鉴定还主要依靠人工，传统的方法是依靠蛾的形态特征进行分类鉴定，这不仅要求专业人员具备一定的专业知识与长期的经验积累，而且鉴定工作耗时，主观性强，因此如何能够利用计算机和有关技术做到对蛾类的自动鉴定已经成为研究人员考虑的课题。IPMist 实验室研发的 BugShape1.0 软件可以提取的数学形态特征有：面积、周长、长轴长、短轴长、等效圆半径、偏心率、紧凑性、球状性、叶状性、圆形性、凸面积、凸周长、外接矩形面积、沿主轴方向外接矩形面积、最小外接矩形面积和不变矩 Hu1、Hu2、Hu3、Hu4、Hu5、Hu6、Hu7，共计 22 项（张建伟，2006）。

自 2010－2011 年利用黑光灯诱集法采集标本，共计收集了 12 科 45 种蛾类昆虫，每种蛾 30 个样本，形成本研究模式的标本数据库。为了从技术角度首先弄清我们拟定的数字化分类方法在蛾类昆虫鉴定中应用的可行性，本文暂不使用该数据库中的全部标本，而先有意识地选出几种加以探讨，并且先从研究右前翅为主，而不是利用昆虫的整个标本特征。这是考虑到：整个标本常会因为展翅程度、触角或足的摆放位置、触角或足的有无等因素对所提取的数学形态特征参数值影响很大，即使同一种类也可能会产生不可避免的误差，蛾翅不仅是蛾等多数昆虫的主要鉴定特征之一，且蛾翅图像便于获取。如果研究证实了本文提出的技术体系是可行的，那么我们将全面扩展应用数字化方法进行蛾类分类鉴定研究；同时，如果研究表明仅以右前翅为主鉴定蛾类昆虫仍有不足，那么我们将利用标本数据库能提供的其他形态特征，逐步提高研究的复杂程度。这样做符合奥卡姆“剃刀原则”，即“如无必要，勿增实体”的“思维经济原则”。

1 材料与方法

1.1 试验材料

本文首次应用特征提取软件 BugShape1.0 并结合图像处理技术，对夜蛾科标本的蛾翅提取与大小尺度和方向均不相关的 13 项数学形态特征数据进行分析，并建立判别准则，探讨蛾翅数学形态特征应用于蛾类分类鉴定的可行性。因此，我们在标本数据库中选择夜蛾科 6 种夜蛾，以它们的右前翅为研究对象，来探索提取的 13 项数学形态特征在蛾类昆虫分类鉴定的可行性。为了对翅型相似的蛾种类也能进行分析，在选择种类时我们选取了夜蛾科地虎属的 3 个种作为试验种类的一部分，而其余 3 种分别属于 3 个属。试验所用夜蛾科种类共有以下 6 种，每个种类取 30 个样本，连同样本采集地点，说明如表 1。

表 1 试验所用 6 种夜蛾名录

Table 1 List of six noctuid species used in the experiment

属名 Genera	种名 Species	采集地 Collecting locality
地虎属 *Agrotis*	小地老虎 *Agrotis ypsilon* (Rottemberg)	河北保定 Baoding, Hebei
	黄地老虎 *Agrotis segetum* (Schiffermiiller)	河北保定 Baoding, Hebei
	八字地老虎 *Agrotis c-nigrum* (Linnaeus)	北京 Beijing
嘴壶夜蛾属 *Oraesia*	鸟嘴壶夜蛾 *Oraesia excavata* Butler	北京 Beijing
铃夜蛾属 *Helicoverpa*	棉铃虫 *Helicoverpa armigera* (Hübner)	河北保定 Baoding, Hebei
裳夜蛾属 *Catocala*	栎光裳夜蛾 *Catocala dissimilis* (Bremer)	北京 Beijing

1.2 图像获取

本研究以蛾翅为基础，主要采用中晶 Phantom v900 Plus 型平板式扫描仪作为主要的图片获取工具。具体数字图像获取步骤如下：

第 1 步，用镊子或手术剪取下完整的昆虫翅膀，翅基不要连带胸部肌肉，以免获取图片时产生不必要的噪点，放在干净纸袋中备用；第 2 步，连接设备，启动扫描程序，设置扫描图片类型为“灰阶”；分辨率为“600DPI”；扫描介质为“负片”；缩放比例为“100%”；第 3 步，进行扫描并保存翅膀图片为 BMP 格式的灰度图片文件，以备分析使用。

1.3 图像处理及特征数据提取

将获取的蛾翅图像进行阈值分割得到二值图像，从二值图像中提取特征数据。由于面积、周长等参数都是绝对量，其值与物体的大小、取像时设备的分辨率、镜头的焦距、物体与摄像设备之间的距离等因素有关，即使本研究采用扫描仪获取图像，这些特征值也与扫描仪的分辨率等参数相关。若采用面积、周长等绝对参数用于识别研究，对获取图像的仪器及其参数设置等都需要严格统一。换句话说，就是在获取图像的仪器及其参数设置等严格统一的条件下，面积和周长等参数用于对昆虫种类的识别是有很重要的作用的。但是，鉴于实际应用中这个条件是比较难于满足的，至少在现阶段，同时也为了日后便于对更多昆虫种类的拓展研究，因此在本研究中只分析与大小尺度和方向均无关的数学形态特征，其中包括矩形度、延长度、叶状性、偏心率、球状性、似圆度和不变矩 Hu1、Hu2 等 13 项特征值(表 2)。以下是对以上 13 项参数的一个简要说明(张建伟, 2006)：

(1)矩形度：物体面积与最小外接矩形面积的比值；

(2)延长度：物体短轴与长轴的比值；

(3)叶状性：边界距质心最短的距离与物体长轴长的比值；

(4)偏心率：由矩特征计算出来的特征值(具体计算方法见附录)；

(5)球状性：目标物体内切圆与外切圆半径的比值；

(6)似圆度：假设目标区域面积为 S，长轴为 L，似圆度为 $4S/\pi L^2$；

(7)~(13)不变矩(Hu)：数字图像的一种统计特征，区域的矩根据所有区域内的点计算得来(具体计算方法见附录)。

1.4 数据处理

试验中每种夜蛾分别用到了 30 个标本，总共有 180 张蛾翅的标本图片，提取了 180 组参数数据。首先采用方差分析对 13 个数学形态特征参数变量进行分析，判断这 13 个参数在不同种之间是否存在差异($P<0.05$)，筛选差异显著的参数，并进行判别分析。

根据方差分析结果对选出的参数进行逐步判别分析，使辨别作用大的指标尽可能不遗漏地进入判别函数，作用小的指标则尽量排除，从而建立较好的判别准则，提高判别效果。错判率的检测用回判法和交叉判别法。

利用逐步判别分析筛选出的特征参数的均值作为聚类分析变量进行聚类分析，对 6 种夜蛾的亲缘关系进行研究。

2 数据分析与结论

2.1 方差分析

本研究利用 SPSS13.0 对 13 项形态特征值进行了多变量方差分析，F 值和 P 值如表 3 所示。结果显示同种不同样本各个参数内部均无显著差异，即

表 2　6 种夜蛾 180 个蛾翅样本的数学形态特征平均值

Table 2　Mean of math-morphological characters (MMCs) extracted from 180 wings of moths of six noctuid species

数学形态特征参数 MMC parameters	八字地老虎 *Agrotis c-nigrum*	小地老虎 *Agrotis ypsilon*	黄地老虎 *Agrotis segetum*	棉铃虫 *Helicoverpa armigera*	鸟嘴壶夜蛾 *Oraesia excavata*	栎光裳夜蛾 *Catocala dissimilis*
矩形度 Rectangularity	0.712	0.698	0.700	0.684	0.706	0.697
延长度 Elongation	0.384	0.369	0.400	0.454	0.507	0.494
叶状性 Lobation	0.395	0.398	0.390	0.373	0.400	0.387
偏心率 Eccentricity	0.584	0.614	0.565	0.466	0.385	0.392
球状性 Sphericity	0.283	0.262	0.279	0.308	0.383	0.361
似圆度 Roundness	0.343	0.325	0.354	0.389	0.455	0.434
Hu1	1.334	1.302	1.359	1.453	1.546	1.562
Hu2	3.375	3.092	3.292	3.671	4.048	4.147
Hu3	6.451	6.049	6.136	6.202	7.043	6.575
Hu4	7.287	7.169	7.472	7.963	9.261	8.706
Hu5	14.003	13.805	14.329	15.200	17.714	16.646
Hu6	9.573	8.833	9.295	10.193	11.851	10.380
Hu7	16.350	15.381	15.513	15.753	17.878	16.774

F 值远远大于 $F(5,174)$。而不同种各个参数间差异极显著（$P<0.01$）。

2.2　判别分析

对 6 种夜蛾的 180 个样本 13 个特征参数进行逐步判别分析，分析结果显示矩形度和延长度两个形态特征对这 6 种夜蛾的分类鉴定没有显著意义，筛选出叶状性、偏心率、球状性、似圆度、Hu1、Hu2、Hu3、Hu4、Hu5、Hu6、Hu7 等 11 个重要指标作为对这 6 种夜蛾进行分类的参数，并且得到标准化判别系数，它们是用来比较各变量对判别值的相对作用程度，哪个变量的标准化系数的绝对值大，就意味着它对判别值有较大影响。利用这些系数和筛选出的 11 个特征参数建立了 5 个标准典型判别函数，用于 6 种夜蛾类昆虫的分类鉴定。它们分别是：

$$Y1 = -0.85a - 0.679b - 0.32c - 0.131d - 0.122e + 0.469f + 0.121g - 0.016h + 0.411i + 0.047j + 0.359k$$

$$Y2 = 0.006a + 0.246b + 0.066c + 0.154d - 0.311e - 0.359f + 0.465g - 0.172h + 0.459i - 0.189j + 0.507k$$

$$Y3 = 0.643a + 0.172b + 1.076c - 0.825d - 0.343e + 0.252f - 0.487g - 0.348h + 0.213i + 0.511j - 0.197k$$

$$Y4 = 0.794a - 0.012b - 1.136c + 0.793d + 0.505e + 0.257f - 0.122g - 0.426h - 0.164i - 0.173j + 0.401k$$

$$Y5 = -0.581a + 0.495b + 0.623c - 1.178d + 0.495e + 0.484f + 0.66g - 0.754h - 0.22i + 0.36j + 0.142k$$

式中：a = 叶状性，b = 偏心率，c = 球状性，d = 似圆度，e = Hu1，f = Hu2，g = Hu3，h = Hu4，i = Hu5，j = Hu6，k = Hu7。

其中前两个函数的累积贡献率已达到 97.5%。由前两个判别函数得出的散点图如图 1 所示，其中横坐标为第一典型判别方程得分，纵坐标为第二典型判别方程得分（以样本数据代入典型判别方程得分计算得到）。从图 1 中可以看出，6 种夜蛾明显地分为了 6 组。同时地虎属的八字地老虎 *Agrotis c-nigrum*、小地老虎 *Agrotis ypsilon* 与黄地老虎 *Agrotis segetum* 3 个种也被很好地区分开，其中八字地老虎可以完全与其他两种地老虎区分开，而小地老虎和黄地老虎则相距较近。

表 4 是判别分析对 180 个样本进行的回归判别和交叉判别结果。结果显示 6 种夜蛾的 180 个样本判别正确率分别为 96.7% 和 93.3%，即在回归判别中只有 1 头小地老虎被误判为黄地老虎，而在交叉判别中有 2 头小地老虎被误判为黄地老虎，其他种类均能够很好地实现分类鉴定。

表 3 方差分析结果

Table 3 Results of variance analysis

数学形态特征参数 MMC parameter	变异来源 Variation source	平方和 Sum of square	自由度 *df*	均方 Mean square	*F* 值 *F* value	*P* 值 *P* value
矩形度 Rectangularity	组间变异 Between groups	0.013	5.000	0.003	12.837	0.000
	组内变异 Within a group	0.036	174.000	0.000		
	总变异 Total	0.049	179.000			
延长度 Elongation	组间变异 Between groups	0.435	5.000	0.087	316.147	0.000
	组内变异 Within a group	0.048	174.000	0.000		
	总变异 Total	0.483	179.000			
叶状性 Lobation	组间变异 Between groups	0.015	5.000	0.003	28.895	0.000
	组内变异 Within a group	0.018	174.000	0.000		
	总变异 Total	0.033	179.000			
偏心率 Eccentricity	组间变异 Between groups	1.369	5.000	0.274	417.729	0.000
	组内变异 Within a group	0.114	174.000	0.001		
	总变异 Total	1.483	179.000			
球状性 Sphericity	组间变异 Between groups	0.332	5.000	0.066	371.363	0.000
	组内变异 Within a group	0.031	174.000	0.000		
	总变异 Total	0.363	179.000			
似圆度 Roundness	组间变异 Between groups	0.347	5.000	0.070	301.395	0.000
	组内变异 Within a group	0.040	174.000	0.000		
	总变异 Total	0.388	179.000			
Hu1	组间变异 Between groups	1.436	5.000	0.287	363.547	0.000
	组内变异 Within a group	0.137	174.000	0.001		
	总变异 Total	1.573	179.000			
Hu2	组间变异 Between groups	23.245	5.000	4.649	418.779	0.000
	组内变异 Within a group	1.932	174.000	0.011		
	总变异 Total	25.177	179.000			
Hu3	组间变异 Between groups	20.184	5.000	4.037	158.031	0.000
	组内变异 Within a group	4.445	174.000	0.026		
	总变异 Total	24.629	179.000			
Hu4	组间变异 Between groups	91.628	5.000	18.326	306.152	0.000
	组内变异 Within a group	10.415	174.000	0.060		
	总变异 Total	102.044	179.000			
Hu5	组间变异 Between groups	323.933	5.000	64.787	277.547	0.000
	组内变异 Within a group	40.616	174.000	0.233		
	总变异 Total	364.549	179.000			
Hu6	组间变异 Between groups	212.481	5.000	42.496	296.411	0.000
	组内变异 Within a group	24.946	174.000	0.143		
	总变异 Total	237.427	179.000			
Hu7	组间变异 Between groups	132.406	5.000	26.481	240.906	0.000
	组内变异 Within a group	19.127	174.000	0.110		
	总变异 Total	151.532	179.000			

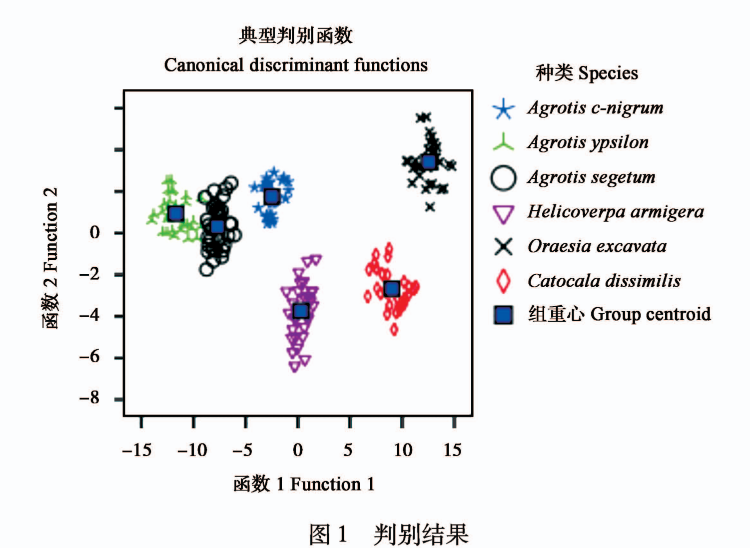


图 1　判别结果

Fig. 1　Result of discriminant analysis

表 4　6 种夜蛾 180 个样本判别分析结果[a,b]

Table 4　Discriminant analysis results[a,b] from 180 specimen of 6 species of moths

		种类 Species	组成员关系预测 Predicted group membership						总和 Total
			Agrotis c-nigrum	*Agrotis ypsilon*	*Agrotis segetum*	*Helicoverpa armigera*	*Oraesia excavata*	*Catocala dissimilis*	
回归 Original	计数 Count	*Agrotis c-nigrum*	30	0	0	0	0	0	30
		Agrotis ypsilon	0	29	1	0	0	0	30
		Agrotis segetum	0	0	30	0	0	0	30
		Helicoverpa armigera	0	0	0	30	0	0	30
		Oraesia excavata	0	0	0	0	30	0	30
		Catocala dissimilis	0	0	0	0	0	30	30
	百分比 Percentage	*Agrotis c-nigrum*	100.0	0	0	0	0	0	100.0
		Agrotis ypsilon	0	96.7	3.3	0	0	0	100.0
		Agrotis segetum	0	0	100.0	0	0	0	100.0
		Helicoverpa armigera	0	0	0	100.0	0	0	100.0
		Oraesia excavata	0	0	0	0	100.0	0	100.0
		Catocala dissimilis	0	0	0	0	0	100.0	100.0
交叉[c] Cross-validated[c]	计数 Count	*Agrotis c-nigrum*	30	0	0	0	0	0	30
		Agrotis ypsilon	0	28	2	0	0	0	30
		Agrotis segetum	0	0	30	0	0	0	30
		Helicoverpa armigera	0	0	0	30	0	0	30
		Oraesia excavata	0	0	0	0	30	0	30
		Catocala dissimilis	0	0	0	0	0	30	30
	百分比 Percentage	*Agrotis c-nigrum*	100.0	0	0	0	0	0	100.0
		Agrotis ypsilon	0	93.3	6.7	0	0	0	100.0
		Agrotis segetum	0	0	100.0	0	0	0	100.0
		Helicoverpa armigera	0	0	0	100.0	0	0	100.0
		Oraesia excavata	0	0	0	0	100.0	0	100.0
		Catocala dissimilis	0	0	0	0	0	100.0	100.0

[a] 已对初始分组案例中的 99.4% 进行了正确分类 99.4% of original grouped cases were correctly classified. [b] 已对交叉验证分组案例中的 98.9% 进行了正确分类 98.9% of cross-validated grouped cases were correctly classified. [c] 仅对分析中的案例进行交叉验证。在交叉验证中，每个案例都是按照从该案例以外的所有其他案例派生的函数来分类的。Cross validation was done only for the cases in the analysis. In cross validation, each case was classified by the functions derived from all cases other than that case.

根据各项特征在 6 种夜蛾分类中所应用的次数，可以计算这些特征的权重，即可以判断该项特征在分类中所起作用的大小。还可以作为将来增加鉴定种类时所采用特征的依据。在分类过程中，一个特征每被使用一次，就加 1（初始值为 0），最后计算每种特征累加值所占所有使用特征累加值之和的百分比。表 5 是各项特征权重的计算结果。

表 5 11 项数学形态特征的权重

Table 5 Weights of 11 math-morphological characters

特征名称 Morphological character	似圆度 Roundness	叶状性 Lobation	偏心率 Eccentricity	球状性 Sphericity	Hu1	Hu2	Hu3	Hu4	Hu5	Hu6	Hu7
权重 Weight（%）	8.82	4.41	16.18	7.35	4.41	14.71	5.88	1.47	16.18	4.41	16.18

2.3 聚类分析

利用逐步判别分析筛选出的 11 个特征参数的均值作为聚类分析变量进行分层聚类分析，结果如图 2 所示。小地老虎和黄地老虎被聚为一类，这与上面的判别分析结果相对应，另外从图中还可看出小地老虎、黄地老虎以及八字地老虎亲缘关系最近，而另外 3 种分属于不同的属，其中铃夜蛾属（棉铃虫 *Helicoverpa armigera*）和地虎属（八字地老虎、小地老虎、黄地老虎）的亲缘关系相对嘴壶夜蛾属（鸟嘴壶夜蛾 *Oraesia excavata*）和裳夜蛾属（栎光裳夜蛾 *Catocala dissimilis*）的亲缘关系较近。在传统昆虫分类学中，小地老虎、黄地老虎以及八字地老虎均属于地虎属，而铃夜蛾属和地虎属均属于切根虫亚科（Agrotinae），裳夜蛾属属于裳夜蛾亚科（Catocalinae），嘴壶夜蛾属属于夜蛾亚科（Noctuinae）。结果表明本研究所得结论与传统昆虫分类学（朱弘复和陈一心，1963；陈一心，1999）相吻合。

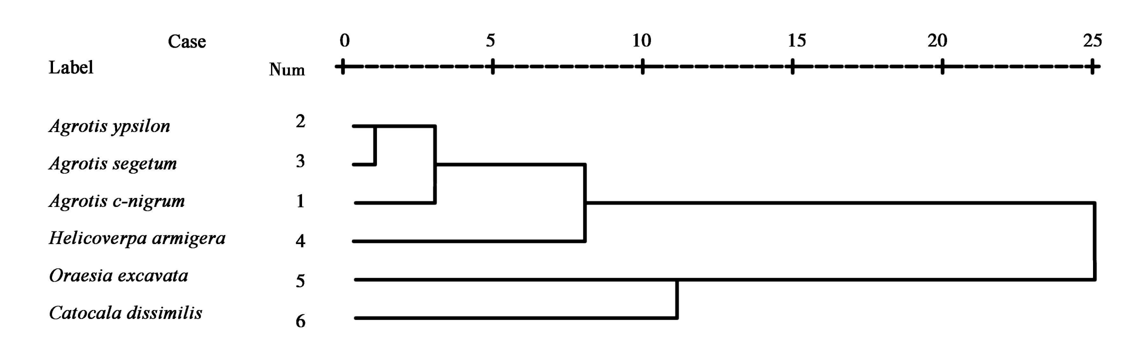


图 2 聚类分析结果

Fig. 2 Result of cluster analysis

2.4 研究结论

通过上述方差分析结果表明所提取的 6 种夜蛾右前翅的 13 个与大小尺度和方向均无关的数学形态特征参数组间差异显著，均可用于分类鉴定。而利用逐步判别分析法筛选出的 11 个判别指标，能够对试验选取的 6 种夜蛾做有效地识别鉴定，这充分证明了利用蛾翅的数学形态特征对夜蛾科分类鉴定的可行性。

筛选出的 11 个数学形态特征在对 6 种夜蛾 180 个样本进行回归判别分析时，其中只有 1 头小地老虎被误判为黄地老虎，其他蛾均被正确地分到独立的组中。从聚类分析中看出小地老虎和黄地老虎二者之间的亲缘关系最近，这可能是造成误判的原因之一，蛾翅轮廓缺损或获取图像轮廓不准确也可能造成误判。这就要求用于提取特征的翅在图像采集时保持其翅面平整，并且保证它的完整性。此外，对于亲缘关系较近的蛾类昆虫还需要探索新的能够进行更准确鉴定的数学形态特征。

11 项数学形态特征值在蛾分类中的相对作用大小依次为：（偏心率、Hu5、Hu7）> Hu2 > 似圆度 > 球状性 > Hu3 >（叶状性、Hu1、Hu6）> Hu4。其中偏心率、Hu5、Hu7 等数学形态特征的权重反映出

它们在分类中的相对作用或者说重要性更大；而 Hu4、叶状性等作用较小。从本文以及其他研究者的文章都共同说明了一点，即数学形态特征在昆虫分类中有很大的应用潜力，但不同的昆虫数学形态特征在不同分类阶元或不同种类上应用的有效性和重要性可能会有所不同。

对 6 种夜蛾亲缘关系的研究所得结论与传统昆虫分类学一致，但是由于本研究的昆虫种类有限，因此只能大概反映出各个种之间亲缘关系的远近。随着研究的深入，研究昆虫种类的增多，本研究所得出的结论可能会得到不断地修正。

3　讨论

本研究使用来自蛾右前翅的数学形态特征即可实现对试验所用夜蛾科 6 种蛾的自动分类鉴定，为未来逐步实现蛾类昆虫的自动识别提供了简便有效的方法。这对于森防站等基层工作人员或者普通林业工作者来说无疑具有重要意义。夜蛾科绝大部分种类均为害虫，本文仅以 6 个种为例探讨了利用数学形态特征对其实现鉴定的可行性，日后可以扩大对夜蛾科种类的研究，但随着需要鉴别蛾的种类的增加，可能需要更多的特征参数才能满足要求，因此剩余的特征参数也将会起到重要的作用，此外，在今后的研究中还可以将蛾后翅、头部、前胸背板、腹部、体躯等部位的数学形态特征应用到昆虫分类中。Weeks 等(1997)曾依据翅脉特征对姬蜂科 5 个姬蜂近似种 175 个样本进行识别研究，准确率达到了 94%。由此可见翅脉的拓扑结构特征对物种的分类识别有着相当重要的作用，这也为后续的研究提供了有力的参考。但是由于夜蛾科所属的鳞翅目其自身的特点，翅面上覆盖大量鳞片，这对于夜蛾科昆虫翅脉特征的研究无疑是一个难点，对去鳞技术与方法也是一种挑战。这需要将多种方法科学合理地结合使用，并不断寻找新的解决方案，以便昆虫翅脉的特征也能应用到夜蛾科或者更多昆虫种类的分类鉴定中。

利用数学形态特征对昆虫进行分类鉴定，需要注意一些问题。对于整个标本常会因为展翅程度、触角或足的摆放位置、触角或足的有无等因素对所提取的数学形态特征参数值影响很大，因此在描述昆虫形态特征时，最好选用一些独立的组织器官，如拆解下来的翅、前胸背板和外生殖器等。在对昆虫形态特征描述的参数选择时，选用与大小尺度和方向均无关的数学形态特征，否则相同的标本会因为放置的方向不一致或取像设备分辨率和镜头的焦距等因素的不一致而导致数据的不同，影响数据的分析与比较。而那些具有绝对值的形态参数如：面积、周长、长轴和短轴等可以在测量个体形态时使用。

相对于本实验利用蛾翅数学形态特征对 6 种夜蛾进行分类的应用研究，传统方法对这 6 种夜蛾的分类鉴定需要用到蛾标本多个部位的形态特征，包括：中足胫节上是否有刺、后翅分岔的不同、额是否有突起、额突的形状、前翅肾纹外侧的黑纵条纹长度等(朱弘复和陈一心，1963)。这些都需要鉴定专家具有相关的专业知识，但是相关专家利用传统的形态特征进行鉴定而累积一定的经验之后，对于有些种类仅通过观察就能迅速准确地作出鉴定，而且对标本没有破坏，保证了标本的完整性和可用性，这对稀有昆虫等标本的鉴定尤为重要。而利用数学形态特征必须经过一些必要的步骤，有些甚至对标本有破坏作用。由此可见，所用的自动鉴定技术既使对蛾的鉴定不受较高专业知识的限制，又解决了目前昆虫鉴定需求不断增多与专业鉴定人员数量、鉴定资源相对减少的矛盾，并且可以减少相关专家大量重复性的鉴定工作。但是如何既能准确地获取昆虫一些独立组织器官的数学形态特征参数又对标本没有破坏作用，这个问题还有待进一步的研究。相信随着获取图像等技术手段的提高和越来越多的昆虫学家研究的深入，这些问题都会迎刃而解。

本实验室赵汗青等(2002，2003) 利用半翅目、鳞翅目和鞘翅目等昆虫整体图像的 11 项数学形态特征参数，研究了昆虫数学形态特征对昆虫实现自动鉴别的可行性；刘芳等(2008)和潘鹏亮等(2008a，2008b)分别利用蝴蝶翅的颜色特征和翅脉几何形态特征实现了对蝴蝶部分种类的自动分类鉴定。本研究蛾翅与大小尺度和方向均不相关的数学形态特征在分类鉴定中的应用也是对这一课题做进一步研究。这些研究不但在鳞翅目昆虫数学形态分类中有重要意义，而且为研究鳞翅目昆虫的系统发育学、物种多样性以及昆虫在不良环境条件下发生变异程度提供了一条便捷途径。比如，目前对昆虫种类系统发育的研究多为分子生物学方法，而利用具有同源性的数学形态特征对昆虫种类的系统发育进行研究，既能够提高效率，又能节省人力、资金等。蛾类昆虫多为农林业的重要害虫，随着杀虫剂

的广泛应用，所带来的昆虫抗药性问题越来越受到人们的关注，而抗药性的强弱或有无在蛾翅的颜色、翅型或翅脉上可能有一定的表现。可以据此评估下一代幼虫的抗药性，从而有针对性地制定更科学的防治策略。随着研究昆虫种类的不断增加和技术手段的提高以及相关专家学者的深入研究，数学形态特征在昆虫分类以及其他领域都会具有越来越好的应用前景。

致谢 衷心感谢中国农业科学院植物保护研究所程登发研究员和张智博士在北京提供的标本采集基地与工具，感谢河北农业大学森保系杨晋宇副教授、苏筱雨实验员、牛晓云和董学超同学以及中国农业大学 IPMist 实验室成员刘钰燕同学在标本采集、标本制作以及图像获取过程中提供的建议和帮助。

参考文献（References）

Chen YX, 1999. Fauna Sinica, Insecta. Vol. 16, Lepidoptera, Ncctuidae. Science Press, Beijing. ［陈一心，1999. 中国动物志：昆虫纲. 第十六卷，鳞翅目：夜蛾科. 北京：科学出版社.］

Hu MK, 1962. Visual pattern recognition by moment invariants. *IEEE Transactions on Information Theory*, 8: 179 – 187.

Lian JB, Wang HQ, 2011. A brief study of the face recognition technology and its application. *Mechanical Management and Development*, (2): 99 – 100. ［连建兵，王浩全，2011. 浅谈人脸识别技术及应用. 机械管理开发，(2)：99 – 100］

Liu F, Shen ZR, Zhang JW, Yang HZ, 2008. Automatic insect identification based on color characters. *Chinese Bulletin of Entomology*, 45(1): 150 – 153. ［刘芳，沈佐锐，张建伟，杨红珍，2008. 基于颜色特征的昆虫自动鉴定方法. 昆虫知识，45(1)：150 – 153］

MacLeod N, 2007. Automated Taxon Identification in Systematics: Theory, Approaches and Applications. CRC Press, London.

O'Neill MA, 2008. DAISY: a practical tool for semi-automated species identification. (URL: http://www.fao.org/ag/agp/agps/c-cab/castudies/pdf/3-001.pdf)

Pan PL, Shen ZR, Yang HZ, Gao LW, Zhang JW, Xie M, 2008a. Digital character extraction and primary analysis of venation of three parnassian butterflies (Lepidoptera, Papilionidae). *Acta Zootaxonomica Sinica*, 33(3): 566 – 571. ［潘鹏亮，沈佐锐，杨红珍，高灵旺，张建伟，谢敏，2008a. 三种绢蝶翅脉数字化特征的提取及初步分析. 动物分类学报，33(3)：566 – 571］

Pan PL, Yang HZ, Shen ZR, Gao LW, Zhang JW, Zhao HQ, Yu XW, 2008b. Research on applying vein feature for mathematical morphology in classification and identification of butterflies (Lepidoptera: Rhopalocera). *Entomotaxonomia*, 30(2): 151 – 160. ［潘鹏亮，杨红珍，沈佐锐，高灵旺，张建伟，赵汗青，于新文，2008b. 翅脉的数学形态特征在蝴蝶分类鉴定中的应用研究. 昆虫分类学报，30(2)：151 – 160］

Prieto C, Munguira M, Romo H, 2009. Morphometric analysis of genitalia and wing pattern elements in the genus *Cupido* (Lepidoptera, Lycaenidae): are *Cupido minimus* and *C. carswelli* different species? *Deul. Entomol. Z.*, 56: 137 – 147.

Wang ZL, 1996. The Development and Application of PQ-INFORMIS – An Information and Identification System for Plant Quarantine Insect Pests. PhD Dissertation, China Agricultural University, Beijing. ［王之岭，1996. 植检害虫图文信息及鉴定辅导系统 PQ-INFORMIS 的研制与应用. 北京：中国农业大学博士学位论文］

Weeks PJD, Gauld ID, Gaston KJ, O'Neill MA, 1997. Automating the identification of insects: a new solution to an old problem. *Bull. Entomol. Res.*, 87: 203 – 211.

Yu XW, 1999. Research and Development of Digital Technology for Insect Imagery. PhD Dissertation, China Agricultural University, Beijing. ［于新文，1999. 昆虫图像数字技术的研究开发. 北京：中国农业大学博士学位论文.］

Zhang FY, Chao BC, Chen XF, 2005. Automatic license plate recognition. *Geospatial Information*, 3(3): 49 – 50. ［张方樱，巢佰崇，陈雪丰，2005. 车牌自动识别系统. 地理空间信息，3(3)：49 – 50］

Zhang JW, 2006. Automatic Identification of Butterflies Based on Computer Vision Technology. PhD Dissertation, China Agricultural University, Beijing. ［张建伟，2006. 基于计算机视觉技术的蝴蝶自动识别研究. 北京：中国农业大学博士学位论文］

Zhao HQ, Shen ZR, Yu XW, 2002. On computer-aided insect identification through math-morphology features. *Journal of China Agricultural University*, 7(3): 38 – 42. ［赵汗青，沈佐锐，于新文，2002. 数学形态特征应用于昆虫自动鉴别的研究. 中国农业大学学报，7(3)：38 – 42］

Zhao HQ, Shen ZR, Yu XW, 2003. Use of math-morphological features in insect taxonomy. I. At the order level. *Acta Entomologica Sinica*, 46(1): 45 – 50. ［赵汗青，沈佐锐，于新文，2003. 数学形态学在昆虫分类学上的应用研究. I. 在目级阶元上的应用研究. 昆虫学报，46(1)：45 – 50］

Zhu HF, Chen YX, 1963. Economic Insect Fauna of China, Fasc. 3. Lepidoptera, Noctuidae (I). Science Press, Beijing. ［朱弘复，陈一心，1963. 中国经济昆虫志. 第三册，鳞翅目：夜蛾科(一). 北京：科学出版社］

附录 Appendix

一、7 个不变矩(Hu)计算方法

矩是数字图像的一种统计特征。区域的矩是用所有属于区域内的点计算出来的。设数字图像中一区域 R 中点 x,y 处的灰度为 $f(x,y)$，则区域的 $p+q$ 阶矩定义为：

$$m_{pq} = \sum_x \sum_y x^p y^q f(x,y) \quad (1\text{-}1)$$

区域 R 的中心矩具有与位置无关的特性，被定义为：

$$\mu_{pq} = \sum_x \sum_y (x-\bar{x})^p (y-\bar{y})^q f(x,y) \quad (1\text{-}2)$$

其中 $\bar{x},\bar{y}$ 相当于区域的重心坐标，被定义为：

$$\bar{x} = m_{10}/m_{00} \quad (1\text{-}3)$$

$$\bar{y} = m_{01}/m_{00} \quad (1\text{-}4)$$

区域 R 的中心矩，按面积规一化后具有与大小尺度无关的特性，规一化后的中心矩被定义为：

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{r}} \text{其中} r = \frac{p+q}{2}+1, p+q = 2,3,\cdots \quad (1\text{-}5)$$

Hu(1962)首次总结了 7 个对平移、旋转和尺度变换不变的距，它是由上述归一化的二阶和三阶中心矩得到的：

$$\phi_1 = \eta_{20} + \eta_{02} \quad (1\text{-}6)$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (1\text{-}7)$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (1\text{-}8)$$

$$\phi_4 = (\eta_{30} - \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (1\text{-}9)$$

$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \mu_{03})^2] \quad (1\text{-}10)$$

$$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \quad (1\text{-}11)$$

$$\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \mu_{03})^2] \quad (1\text{-}12)$$

二、偏心率计算方法

由以上矩特征可以按如下公式计算物体的偏心率：

$$\text{Eccentricity} = \frac{(\mu_{02} - \mu_{20})^2 + 4\mu_{11}}{area} \quad (1\text{-}13)$$

(责任编辑：袁德成)